Technical Note

Sampling Variability

The tables in *Earnings and Employment Data for Wage and Salary Workers Covered Under Social Security, by State and County, 1993*, present data that are from a 1-percent sample file drawn from the Social Security Administration's administrative records.

Because of sampling variability, estimates based on sample data may differ from the figures that would have been obtained had all rather than specified samples of the records been used. The standard error is a measure of sampling variability. About 68 percent of all possible probability samples selected with the same specifications will give estimates within one standard error of the figure obtained from the compilation of all records. Similarly, about 95 percent will give estimates within two standard errors, and about 99 percent will give estimates within two and one-half standard errors. The standard error of an estimate depends on the design elements, such as the method of sampling, sample size, and the estimation process.

Because of the large number of data cells tabulated from the sample files, it is not practical to calculate the standard error for every possible cell. However, standard errors for a large number of cells were estimated. These estimates were used to fit regression curves to provide estimates of approximate standard errors associated with tabulated counts and proportions.

The tables that follow show the sampling variability and provide a general order of magnitude for similar estimates from the various sample files. Table A presents approximate standard errors for the estimated number of persons from the 1-percent sample file. The reliability of an estimated percentage depends on both the size of the percentage and on the size of the total on which the percentage is based. Data in table B provide approximations of the standard errors of the estimated percentage of persons in the 1-percent sample file. The standard errors are expressed in percentage points, and the bases shown are expressed in terms of the estimated total population.

Table A.—Approximations of standard errors for estimated number of persons

1-percent file	
Size of estimate (inflated)	Standard error
500	250
1,000	300
2,500	500
5,000	800
7,500	900
10,000	1,100
25,000	1,700
50,000	2,400
75,000	3,000
100,000	3,400
250,000	5,400
500,000	7,800
750,000	9,600
1,000,000	11,100
5,000,000	25,800
10,000,000	36,900
25,000,000	57,700
50,000,000	76,100
75,000,000	82,900

Table B.—Approximations of standard errors for estimated percentage of persons from 1-percent sample file

Size of base (inflated)	Estimated percentage				
	2 or 98	5 or 95	10 or 90	25 or 75	50
1,000	4.7	7.3	10.1	14.5	16.8
10,000	1.5	2.3	3.2	4.6	5.3
50,000	.7	1.0	1.4	2.1	2.4
100,000	.5	.7	1.0	1.5	1.7
1,000,000	.1	.2	.3	.5	.5
5,000,000	.1	.1	.1	.2	.2
10,000,000	([1])	.1	.1	.2	.2
50,000,000	([1])	([1])	([1])	.1	.1
100,000,000	([1])	([1])	([1])	([1])	([1])

[1] Less than 0.05 percent.